



University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): M Hairer and NS Pillai

Article Title: Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion

Year of publication: 2009

Link to published article:

<http://www2.warwick.ac.uk/fac/sci/statistics/crism/research/2009/paper09-38>

Publisher statement: None

Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion

September 28, 2009

M. Hairer^{1,2}, N.S. Pillai³

¹ Mathematics Institute, The University of Warwick
Email: M.Hairer@Warwick.ac.uk

² Courant Institute, New York University
Email: mhairer@cims.nyu.edu

³ CRiSM, The University of Warwick
Email: n.pillai@warwick.ac.uk

Abstract

We demonstrate that stochastic differential equations (SDEs) driven by fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$ have similar ergodic properties as SDEs driven by standard Brownian motion. The focus in this article is on hypoelliptic systems satisfying Hörmander's condition. We show that such systems satisfy a suitable version of the strong Feller property and we conclude that they admit a unique stationary solution that is physical in the sense that it does not “look into the future”.

The main technical result required for the analysis is a bound on the moments of the inverse of the Malliavin covariance matrix, conditional on the past of the driving noise.

Keywords: Ergodicity, Fractional Brownian motion, Hörmander's theorem

1 Introduction

In this paper we study the ergodic properties of stochastic differential equations (SDEs) of the type

$$dx = V_0(x) dt + \sum_{i=1}^d V_i(x) \circ dB^i(t), \quad x(0) \in \mathbb{R}^n, \quad (1.1)$$

driven by independent fractional Brownian motions (fBm) B^i with fixed Hurst index $H \in (1/2, 1)$. Recall that fBm is the centred Gaussian process with $B^i(0) = 0$ and

$$\mathbb{E}|B^i(t) - B^i(s)|^2 = |t - s|^{2H}.$$

Although the basic theory of SDEs driven by fBm (for $H \in (1/4, 1)$) is now well established (see for example [Nua06, Mis08, CQ02, Fri10]), the ergodicity of such SDEs has only been studied recently. The main difficulty is that non-overlapping increments of fBm are not independent, so that we are dealing with processes that do not have the Markov property. As a consequence, the traditional ergodic theory for Markov processes as in [MT93] does not apply to this situation.

To our knowledge, the first result on the ergodicity of (1.1) was given in [Hai05], where the author considers the case where the vector fields V_i for $i > 0$ are *constant* and *non-degenerate* in the sense that the corresponding matrix is of rank n . In this case, an explicit coupling argument allows to show that any solution converges to the unique stationary solution at speed bounded above by approximately $t^{-1/8}$. The argument however relies in an essential way on the additivity of the noise.

A more general construction was given in [HO07, Hai09], where the authors built a general framework of stochastic dynamical systems (SDSs) very similar to that of random dynamical systems (RDSs) [Arn98]. In this framework, the notion of an “invariant measure” can be defined for an autonomous equation driven by an ergodic process with stationary but not necessarily independent increments. Besides technical questions of continuity, the main difference between the framework of SDSs and that of RDSs is one of viewpoint. An RDS is considered as a dynamical system on the state space times the space of all ‘futures’ of the driving process, endowed with a cocycle structure. An SDS on the other hand is considered as a Markov process on the state space times the space of all ‘pasts’ of the driving process, endowed with a cocycle structure. This shift of viewpoint has two important advantages when studying the ergodic properties of such systems:

- Since the ‘future’ of the driving noise is not part of the augmented phase space, we do not have to deal with ‘unphysical’ stationary solutions, where the value of the process at a given time is a function of the future realisation of the driving noise.
- It allows to build our intuition by taking as basis the theory of Markov processes, rather than the theory of dynamical systems. There, many criteria for the uniqueness of an invariant measure are known to exist.

Inspired by the celebrated Doob-Khasminskii theorem (see for example [DPZ96]), [HO07, Hai09] introduced a notion of a strong Feller property for SDSs that is a natural generalisation of the same notion for Markov semigroups. It turns out that for a large class of ‘quasi-Markovian’ SDSs (this includes SDEs of the type (1.1), but also many other examples), uniqueness of the invariant measure then follows like in the Markovian case from the strong Feller property, combined with some type of topological irreducibility. To illustrate these results, denote by $A_{t,x}$ the closure of the set of all points that are accessible at time t for solutions to (1.1) starting at x for some smooth control B . Then, one has

Theorem 1.1 *If there exists $t > 0$ such that (1.1) is strong Feller at time t and such that $\bigcap_{x \in \mathbb{R}^n} A_{t,x} \neq \emptyset$, then (1.1) can have at most one invariant measure.*

Remark 1.2 Since (1.1) does not determine a Markov process, it is not clear *a priori* what is the correct notion of an invariant measure. The notion retained here is the one introduced in [HO07, Hai09]. Essentially, an invariant measure for (1.1) is the law of a stationary solution which does not ‘look into the future’. Similarly, it is not clear what is the correct notion of a ‘strong Feller property’ for such a system. Precise definitions can be found in Section 2.1 below.

Furthermore, it was shown in [HO07] that if (1.1) is ‘elliptic’ in the sense that the matrix determined by $\{V_i(x)\}_{i>0}$ is of full rank at every $x \in \mathbb{R}^n$, then it does indeed satisfy the strong Feller property. In the case of SDEs driven by white noise, it is known on the other hand that the strong Feller property holds whenever the vector

fields V_i satisfy Hörmander's bracket condition, namely that the Lie algebra generated by $\{\partial_t + V_0, V_1, \dots, V_d\}$ spans \mathbb{R}^{n+1} (the additional coordinate corresponds to the ∂_t direction) at every point. See also Assumption 4.1 below for a precise formulation.

Note at this stage that the strong Feller property of (1.1) is closely related to the existence of densities for the laws of its solutions, but cannot be deduced from it directly. The existence of smooth densities under an ellipticity assumption was shown in [HN07, NS06], while it was shown under the assumption of Hörmander's bracket condition in [BH07]. In this paper, we are going to address the question of showing that the strong Feller property (in the sense given in [HO07, Hai09], see also Definition 2.1 below) holds for (1.1) under Hörmander's bracket condition.

Our main result is:

Theorem 1.3 *Assume that the vector fields $\{V_i\}$ have bounded derivatives of all orders for $i \geq 0$ and are furthermore bounded for $i \geq 1$. Then, if Hörmander's bracket condition holds at every point, (1.1) is strong Feller. In particular, if there exists $t > 0$ such that $\bigcap_{x \in \mathbb{R}^n} A_{t,x} \neq \emptyset$ and if there exist constants $c, C > 0$ such that*

$$\langle x, V_0(x) \rangle \leq C - c|x|^2, \quad (1.2)$$

then (1.1) admits exactly one invariant measure.

Remark 1.4 Since we would like to satisfy (1.2) in order to guarantee the existence of an invariant measure by [HO07, Prop 4.6], we do not impose that V_0 is bounded. This creates some technical difficulties in obtaining a priori bounds on the solutions which are not present in [BH07].

Remark 1.5 The notion of an 'invariant measure' studied in [BC07] is much more restrictive than the one studied here, since they considered the situation where the initial condition of the equations is *independent* of the driving noise, whereas we consider the case where interdependence with the past of the driving noise is allowed. In particular, even the case $n = d = 1$ with $V_0(x) = -x$ and $V_1(x) = 1$ admits no invariant measure in the sense of [BC07].

On the other hand, our definition of an invariant measure is more restrictive than that of a random invariant measure in [Arn98]. Since it generalises the notion of an invariant measure for a Markov process, this is to be expected. Indeed, there are simple examples of elliptic diffusions on the circle whose Markov semigroup admits a unique invariant measure, but that admit more than one random invariant measure when viewed as RDS's.

The remainder of the article is structured as follows. In Section 2, we set up our notations and recall the relevant results from [Hai05] and [HO07]. In Section 3, we derive the necessary moment estimates for the solutions of the SDE (1.1). In Section 4, we then obtain an invertibility result on the conditioned Malliavin covariance matrix of the process (1.1), provided that Hörmander's condition holds. Similarly to [BH07], this allows us to show the smoothness of the laws of solutions to (1.1), conditional on the past of the driving noise. In Section 5, we finally show that the strong Feller property is satisfied for (1.1) under Hörmander's condition, thereby concluding the proof of Theorem 1.3.

2 Preliminaries

In this section, we describe the general framework in which we view solutions to (1.1) and we recall some of the basic results from [Hai05, HO07, Hai09].

2.1 General framework

For $\gamma, \delta \in (0, 1)$ define $\mathcal{W}_{(\gamma, \delta)}$ to be the completion of $\mathcal{C}_0^\infty(\mathbb{R}_-; \mathbb{R}^d)$ with respect to the norm

$$\|\omega\|_{(\gamma, \delta)} \equiv \sup_{t, s \in \mathbb{R}_-} \frac{|\omega(t) - \omega(s)|}{|t - s|^\gamma (1 + |t| + |s|)^\delta} . \quad (2.1)$$

We write $\widetilde{\mathcal{W}}_{(\gamma, \delta)}$ for the corresponding space containing functions on \mathbb{R}_+ instead of \mathbb{R}_- . Note that when restricted to a compact time interval, the norm (2.1) is equivalent to the usual Hölder norm with exponent γ . Moreover, $\mathcal{W}_{(\gamma, \delta)}$ is a separable Banach space.

For $H \in (1/2, 1)$, $\gamma \in (1/2, H)$ and $\gamma + \delta \in (H, 1)$, it can be shown that there exists a Borel probability measure \mathbb{P}_H on $\Omega = \mathcal{W}_{(\gamma, \delta)} \times \widetilde{\mathcal{W}}_{(\gamma, \delta)}$ such that the canonical process associated to \mathbb{P}_H is a two-sided fractional Brownian motion with Hurst parameter H [Hai05, HO07]. We fix such values γ , δ , and H once and for all and we drop the subscripts (γ, δ) from the spaces \mathcal{W} and $\widetilde{\mathcal{W}}$ from now on. Note that \mathbb{P}_H is a product measure if and only if $H = 1/2$, but in general we can disintegrate it into a projected measure (denoted again by \mathbb{P}_H) on \mathcal{W} and regular conditional probabilities $\mathcal{P}(\omega, \cdot)$ on $\widetilde{\mathcal{W}}$. Since \mathbb{P}_H is Gaussian, there is a unique choice of $\mathcal{P}(\omega, \cdot)$ which is weakly continuous in ω .

We denote by $\varphi: \mathbb{R}^n \times \widetilde{\mathcal{W}} \rightarrow \mathcal{C}(\mathbb{R}_+, \mathbb{R}^n)$ the solution map to (1.1). It is well-known [You36, HN07] that if the vector fields V_i have bounded derivatives, then this solution is well-defined in a pathwise sense (integrals are simply Riemann-Stieltjes integrals) for all times. We also define the shift map $\theta_t: \mathcal{W} \times \widetilde{\mathcal{W}} \rightarrow \mathcal{W} \times \widetilde{\mathcal{W}}$ by identifying elements $(\omega, \tilde{\omega}) \in \mathcal{W} \times \widetilde{\mathcal{W}}$ with the corresponding ‘concatenated’ function $(\omega \sqcup \tilde{\omega}): \mathbb{R} \rightarrow \mathbb{R}^d$ and setting

$$\theta_t(\omega, \tilde{\omega}) \sim (\omega \sqcup \tilde{\omega})(t + s) - (\omega \sqcup \tilde{\omega})(s) .$$

Denote $\varphi_t(x, \omega) = \varphi(x, \omega)(t)$ and denote by $\Phi_t: \mathbb{R}^n \times \mathcal{W} \times \widetilde{\mathcal{W}} \rightarrow \mathbb{R}^n \times \mathcal{W}$ the augmented solution map given by

$$\Phi_t(x, \omega, \tilde{\omega}) = (\varphi_t(x, \tilde{\omega}), \Pi_{\mathcal{W}}\theta_t(\omega, \tilde{\omega})) ,$$

where $\Pi_{\mathcal{W}}$ is the projection onto the \mathcal{W} -component. With this notation, we can view the solutions to (1.1) as a Markov process on $\mathbb{R}^n \times \mathcal{W}$ with transition probabilities given by

$$\mathcal{Q}_t(x, \omega; \cdot) = \Phi_t(x, \omega, \cdot)^* \mathcal{P}(\omega, \cdot) .$$

These transition probabilities can actually be shown to be Feller [HO07], but they are certainly *not* strong Feller in the usual sense, since transition probabilities starting from different instances of ω remain mutually singular for all times. Instead, we will use a notion of strong Feller property that is better suited for the particular structure of the problem at hand, see Definition 2.1 below.

However, the question of uniqueness of the invariant measure for (1.1) should not be interpreted as the question of uniqueness of the invariant measure for \mathcal{Q}_t . This is because one might imagine that the augmented phase space $\mathbb{R}^n \times \mathcal{W}$ contains some ‘redundant’ randomness that is not used to describe the stationary solutions to (1.1).

(This would be the case for example if the V_i 's are not always linearly independent.) One would like therefore to have a concept of uniqueness for the invariant measure that is independent of the particular description of the driving noise.

To this end, we introduce the Markov transition kernel \bar{Q} from $\mathbb{R}^n \times \mathcal{W}$ to $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^n)$ given by

$$\bar{Q}(x, \omega; \cdot) = \varphi(x, \cdot)^* \mathcal{P}(\omega, \cdot) .$$

This is the conditional law of the solution to (1.1) given a realisation ω of the past of the driving noise. This now defines a natural equivalence relation on measures on $\mathbb{R}^n \times \mathcal{W}$ by

$$\mu \sim \nu \quad \Leftrightarrow \quad \bar{Q}\mu = \bar{Q}\nu .$$

With these definitions at hand, the statement ‘the invariant measure for (1.1) is unique’ should be interpreted as ‘the Markov semigroup Q_t has a unique invariant measure, modulo the equivalence relation \sim ’.

2.2 Ergodicity of SDEs driven by fBm

We now summarise some of the relevant results from [Hai05, HO07, Hai09] giving conditions for the uniqueness of the invariant measure for (1.1). This requires a notion of ‘strong Feller’ property for (1.1). We stress again that the definition given here has *nothing to do* with the strong Feller property of Q_t . It rather generalises the notion of the strong Feller property for the Markov process associated to (1.1) in the case where the driving noise is white in time.

Let $R_t: \mathcal{C}(\mathbb{R}_+, \mathcal{X}) \mapsto \mathcal{C}([t, \infty), \mathcal{X})$ denote the natural restriction map and let $\|\cdot\|_{TV}$ denote the total variation norm. We then say that:

Definition 2.1 The solutions to (1.1) are said to be *strong Feller* at time t if there exists a jointly continuous function $\ell: (\mathbb{R}^n)^2 \times \mathcal{W} \rightarrow \mathbb{R}_+$ such that

$$\|R_t^* \bar{Q}(x, \omega; \cdot) - R_t^* \bar{Q}(y, \omega; \cdot)\|_{TV} \leq \ell(x, y, \omega) , \quad (2.2)$$

and $\ell(x, x, \omega) = 0$ for every $x \in \mathbb{R}^n$ and every $\omega \in \mathcal{W}$.

We also introduce the following notion of irreducibility:

Definition 2.2 The solutions to (1.1) are said to be *topologically irreducible* at time t if, for every $x \in \mathbb{R}^n$, $\omega \in \mathcal{W}$ and every non-empty open set $U \subset \mathbb{R}^n$, one has $Q_t(x, \omega; U \times \mathcal{W}) > 0$.

The following result from [HO07, Thm 3.10] is the main abstract uniqueness result on which this article is based:

Theorem 2.3 *If there exist times $s > 0$ and $t > 0$ such that the solutions to (1.1) are strong Feller at time t and irreducible at time s , then (1.1) can have at most one invariant measure.*

Therefore, in order to prove Theorem 1.3, the only missing ingredient that we need to establish is the strong Feller property for (1.1) under Hörmander’s bracket condition.

2.3 Notations and definitions

For $T > 0$ and measurable $f : [0, T] \mapsto \mathbb{R}^n$, set

$$\|f\|_{0,T,\infty} = \sup_{t \in [0,T]} |f(t)|, \quad \|f\|_{0,T,\gamma} = \sup_{s,t \in [0,T]} \frac{|f(t) - f(s)|}{|t - s|^\gamma}. \quad (2.3)$$

For $\alpha \in (0, 1)$, we also define the fractional integration operator \mathcal{I}^α and the corresponding fractional differentiation operator \mathcal{D}^α by

$$\begin{aligned} \mathcal{I}^\alpha f(t) &\equiv \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds, \\ \mathcal{D}^\alpha f(t) &\equiv \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t (t-s)^{-\alpha} f(s) ds. \end{aligned} \quad (2.4)$$

Remark 2.4 The operators \mathcal{I}^α and \mathcal{D}^α are inverses of each other, see [SKM93] for a survey of fractional integral operators.

The reason why these operators are crucial for our analysis is that the Markov transition kernel \mathcal{P} appearing in Section 2.1 is given by translates of the image of Wiener measure under $\mathcal{I}^{H-1/2}$, see also Lemma 2.7 below.

Indeed, by the celebrated Mandelbrot-Van Ness representation of the fBm [MVN68], we may express the two-sided fBm B with Hurst parameter $H \in (0, 1)$ in terms of a two-sided standard Brownian motion W as

$$B_t = \alpha_H \int_{-\infty}^0 (-r)^{H-1/2} (dW_{r+t} - dW_r) \quad (2.5)$$

for some $\alpha_H > 0$. The advantage of this representation is that it is invariant under time-shifts, so that it is natural for the study of ergodic properties, see [ST94] for more details.

Define now the operator $\mathcal{G}: \mathcal{W} \rightarrow \widetilde{\mathcal{W}}$ by:

$$\mathcal{G}\omega(t) \equiv \gamma_H \int_0^\infty \frac{1}{r} g\left(\frac{t}{r}\right) \omega(-r) dr, \quad (2.6)$$

where the kernel g is given by

$$g(x) \equiv x^{H-1/2} + (H-3/2)x \int_0^1 \frac{(u+x)^{H-5/2}}{(1-u)^{H-1/2}} du, \quad (2.7)$$

and the constant γ_H is given by $\gamma_H = (H-1/2)\alpha_H\alpha_{1-H}$. Here, α_H is the constant appearing in (2.5). It was shown in [HO07, Prop A.2] that \mathcal{G} is indeed a bounded linear operator from \mathcal{W} into $\widetilde{\mathcal{W}}$. Furthermore, we can quantify its behaviour near $t = 0$ in the following way:

Lemma 2.5 *On any time interval bounded away from 0, the map $t \mapsto \mathcal{G}\omega(t)$ is \mathcal{C}^∞ . Furthermore, if we set $f_\omega(t) = t \frac{d}{dt} \mathcal{G}\omega(t)$, then we have $f_\omega(0) = 0$ and, for every $T > 0$, there exists a constant M_T such that $\|f_\omega\|_{0,T,\gamma} < M_T \|\omega\|_{(\gamma,\delta)}$.*

The proof of this result is postponed to the appendix.

Remark 2.6 Since \mathbb{P}_H is Gaussian, it follows in particular that both $\|\mathcal{G}\omega\|_{0,T,\gamma}$ and $\|f_\omega\|_{0,T,\gamma}$ have exponential moments under \mathbb{P}_H by Fernique's theorem [Bog98].

Let $\tau_h: w \mapsto w + h$ denote the translation map on $\widetilde{\mathcal{W}}$. We cite the following result from [HO07, Lemma 4.2]:

Lemma 2.7 *The regular conditional probabilities $\mathcal{P}(\omega, \cdot)$ of \mathbb{P}_H given $\omega \in \mathcal{W}$ are*

$$\mathcal{P}(w, \cdot) = (\tau_{\mathcal{G}w} \circ \mathcal{I}^{H-\frac{1}{2}})^* \mathbf{W},$$

where \mathbf{W} is the standard Wiener measure over \mathbb{R}_+ and \mathcal{I}^α is as defined in (2.4).

We will henceforth interpret the above lemma in the following way. The driving fractional Brownian motion B can be written as the sum of two independent processes:

$$B_t \stackrel{d}{=} \tilde{B}_t + (\mathcal{G}\omega)(t) \stackrel{d}{=} \tilde{B}_t + m_t, \quad \tilde{B}_t = \alpha_H \int_0^t (t-s)^{H-1/2} dW(s), \quad (2.8)$$

where W is a standard Wiener process independent of the ‘past’ $\omega \in \mathcal{W}$. This notation will be used repeatedly in the sequel. We will also repeatedly use the notations $\tilde{\mathbb{E}}$ and $\tilde{\mathbb{P}}$ for expectations and probabilities over $\tilde{\omega}$ with ω fixed.

3 Estimates on the solutions

In this section, we derive moment estimates for the solutions to equations of the form:

$$X_t^x = V_0(X_t) dt + \sum_{i=1}^d V_i(X_t) \circ dB_t^i, \quad t \geq 0, \quad (3.1)$$

where $x = X_0 \in \mathbb{R}^n$ and $V_i, i = 1, 2, \dots, d$ are bounded, C^∞ vector fields in \mathbb{R}^n with bounded derivatives, and V_0 is a possibly *unbounded* C^∞ vector field with bounded derivatives. Our bounds will be purely pathwise bounds, so the fact that the B^i ’s are sample paths of fractional Brownian motion is irrelevant, except to get a bound on their Hölder regularity.

All that we will assume is that the driving process is γ -Hölder continuous for some $\gamma > 1/2$. The main result in this section are a priori bounds not only on (3.1), but also on its Jacobian and its second variation with respect to the initial condition. This will be a slight generalisation of the results in [HN07], which didn’t allow for the drift term V_0 to grow.

Let $\Phi(t, x) \equiv X_t^x$ denote the flow map solving (3.1) and define its Jacobian by

$$J_{0,t} \equiv \frac{\partial \Phi(t, x)}{\partial x}.$$

For notational convenience, set $V = (V_1, V_2, \dots, V_d)$. Then, we can write (3.1) in compact form:

$$dX_t = V_0(X_t) dt + V(X_t) \circ dB_t, \quad (3.2a)$$

$$dJ_{0,t} = V'_0(X_t) J_{0,t} dt + V'(X_t) J_{0,t} \circ dB_t, \quad (3.2b)$$

$$dJ_{0,t}^{-1} = -J_{0,t}^{-1} V'_0(X_t) dt - J_{0,t}^{-1} V'(X_t) \circ dB_t. \quad (3.2c)$$

Here, both J and J^{-1} are $n \times n$ matrices, and $J_{0,0} = J_{0,0}^{-1} = 1$. One crucial ingredient in order to obtain the bounds on the Malliavin matrix required to show the strong Feller property are control on the moments of both the solution and its Jacobian. These bounds do not quite follow from standard results since most of these assume that V_0 is also bounded, which we cannot assume since we need (1.2) in order to have a non-empty statement.

3.1 Moments estimates

Let $\mathcal{L}_k(\mathcal{X}_1, \dots, \mathcal{X}_k; \mathcal{Y})$ denote the set of k -multilinear maps from $\mathcal{X}_1 \times \dots \times \mathcal{X}_k$ to \mathcal{Y} . As usual, \mathcal{L}_1 is denoted by \mathcal{L} . In this section, we consider a system of differential equations of the form

$$x_t = x_0 + \int_0^t f_0(x_r) dr + \int_0^t f(x_r) dB_r, \quad (3.3a)$$

$$y_t = y_0 + \int_0^t g(x_r)(y_r, dB_r), \quad (3.3b)$$

where $x_t \in \mathbb{R}^n$, $y_t \in \mathbb{R}^m$, $B : [0, \infty) \mapsto \mathbb{R}^d$ is a Hölder function of order $\gamma > 1/2$ and $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f : \mathbb{R}^n \mapsto \mathcal{L}(\mathbb{R}^d; \mathbb{R}^n)$, $g : \mathbb{R}^n \mapsto \mathcal{L}_2(\mathbb{R}^m, \mathbb{R}^d; \mathbb{R}^m)$ are given \mathcal{C}^1 functions. Note that (3.2) is indeed of this form with $m = 2n^2$.

Remark 3.1 The reason why we don't include a ' dr ' term in (3.3b) is because this is already covered by the present result by setting $B_1(r) = r$ for example. Treating this term separately like in (3.3a) might allow us to slightly improve our results, but the present formulation is sufficient for our purpose.

Regularity of solutions to equations of this kind has been well studied, pioneered by the technique of Young ([You36]) and more recently using the theory of Rough paths (see for example [LQ02, Fri10] and the references therein). Using fractional derivatives, it was shown in [HN07, Thm 3] that

$$\sup_{0 \leq t \leq T} |y_t| \leq 2^{1+M} \|B\|_{0,T,\gamma}^{1/\gamma} |y_0|, \quad (3.4)$$

where $\|B\|_{0,T,\gamma}$ is the Hölder norm defined in (2.3) and M is a constant depending on the supremum norms of f_0 , f , g and their first derivatives. Unfortunately, these estimates are not sufficient to obtain moment bounds on the Jacobian $J_{0,t}$ and its inverse, since we wish to consider situations where f_0 is unbounded, so that its supremum norm is not finite.

However it turns out that an estimate of the type (3.4) holds even if only the derivative of f_0 is bounded, thanks to the fact that the corresponding "driving noise" (t) is actually a differentiable function. This is the content of the following result:

Lemma 3.2 *For the process y_t defined in (3.3b), we have the pathwise bounds:*

$$\|x\|_{0,T,\gamma} \leq M(1 + |x_0|) (1 + \|B\|_{0,T,\gamma})^{1/\gamma} \quad (3.5a)$$

$$\|y\|_{0,T,\gamma} \leq M(1 + |x_0|) |y_0| e^{M\|B\|_{0,T,\gamma}^{1/\gamma}} \quad (3.5b)$$

where M is an absolute constant which depends only on $\|f\|_\infty$, $\|f'\|_\infty$, $\|f'_0\|_\infty$, $\|g\|_\infty$, $\|g'\|_\infty$ and T .

Proof. The proof is almost identical to that of [HN07], so we only try to highlight the main differences. For $s, t \in [0, T]$, we have the apriori bound (see [HN07, p. 403]),

$$\left| \int_s^t f(x_r) dB_r \right| \leq M\|B\|_{0,T,\gamma} \left((t-s)^\gamma + \|x\|_{s,t,\gamma} (t-s)^{2\gamma} \right),$$

where the constant M is independent of $\|B\|_{0,T,\gamma}$, but depends on f . Since f_0 grows almost linearly, we also have

$$\left| \int_s^t f_0(x_r) dr \right| \leq M \left((t-s) + |x_s|(t-s) + \|x\|_{s,t,\gamma}(t-s)^{\gamma+1} \right),$$

and therefore

$$\sup_{s \leq m, n \leq t} \left| \int_m^n f_0(x_r) dr \right| \leq M \left((t-s) + |x_s|(t-s) + 2\|x\|_{s,t,\gamma}(t-s)^{\gamma+1} \right). \quad (3.6)$$

Thus we have

$$\|x\|_{s,t,\gamma} \leq M_1 \left(1 + \|B\|_{0,T,\gamma} + |x_s|(t-s)^{1-\gamma} + (\|B\|_{0,T,\gamma} + 1)\|x\|_{s,t,\gamma}(t-s)^\gamma \right)$$

for a constant M_1 depending on f and the terminal time T . Set $\Delta = (2M_1(\|B\|_{0,T,\gamma} + 1))^{-1/\gamma}$. Then for $|t-s| < \Delta$,

$$\|x\|_{s,t,\gamma} \leq 2M_1 \left(1 + \|B\|_{0,T,\gamma} + |x_s|\Delta^{1-\gamma} \right). \quad (3.7)$$

Since by definition we have:

$$\|x\|_{s,t,\infty} \leq |x_s| + \|x\|_{s,t,\gamma}|t-s|^\gamma,$$

for $|t-s| < \Delta$, from (3.7) it follows that

$$\|x\|_{s,t,\infty} \leq |x_s|(1 + 2M_1\Delta) + 1. \quad (3.8)$$

Iterating the above estimate for $N = T/\Delta$, it follows that

$$\begin{aligned} \|x\|_{0,T,\infty} &\leq |x_0|(1 + 2M_1\Delta)^N + \sum_{k=1}^{N-1} (1 + 2M_1\Delta)^k \leq (1 + |x_0|)(1 + 2M_1\Delta)^N \\ &= (1 + |x_0|)(1 + 2M_1\Delta)^{T/\Delta}. \end{aligned}$$

Using the fact that $(1 + \frac{x}{\Delta})^\Delta \leq e^x$ for $\Delta > 0$, this finally yields

$$\|x\|_{0,T,\infty} \leq (1 + |x_0|) e^{2M_1 T} T (2M_1(\|B\|_{0,T,\gamma} + 1))^{1/\gamma}. \quad (3.9)$$

Remark 3.3 In the last step of the above argument we bounded above $(1 + 2M_1\Delta)^{T/\Delta}$ by $e^{2M_1 T}$ and thus obtained a constant which is exponential in T , but *independent* of $\Delta = \mathcal{O}(\|B\|_{0,T,\gamma}^{-1/\gamma})$. If the noise corresponding to the vector field f_0 were *not* differentiable but only Hölder continuous, then instead of (3.8), then the exponent would include a power of the Hölder constant of the driving noise, thus yielding an estimate comparable to those obtained in [HN07].

Thus substituting the bound (3.9) in (3.7) yields that, for $t-s < \Delta$,

$$\|x\|_{s,t,\gamma} \leq M(1 + |x_0|)(1 + \|B\|_{0,T,\gamma}).$$

The subadditivity of the Hölder norm then immediately implies the bound

$$\|x\|_{0,T,\gamma} \leq M(1 + |x_0|)(1 + \|B\|_{0,T,\gamma})^{1-\gamma} = M(1 + |x_0|)(1 + \|B\|_{0,T,\gamma})^{1/\gamma},$$

proving the claim made in (3.5a).

Now we come to the second part of Lemma 3.2. Once again, for $s, t \in [0, T]$ we have the apriori bound (see [HN07, p. 406])

$$\left| \int_s^t g(x_r) y_r dB_r \right| \leq \tilde{M}(1 + \|B\|_{0,T,\gamma}) \left(\|g\|_\infty \|y\|_{s,t,\infty} (t-s)^\gamma + (\|g\|_\infty \|y\|_{s,t,\gamma} + \|g'\|_\infty \|y\|_{s,t,\infty} \|x\|_{s,t,\gamma}) (t-s)^{2\gamma} \right) \quad (3.10)$$

for a constant \tilde{M} independent of B . Thus, setting as before $\Delta = (2M_1(\|B\|_{0,T,\gamma} + 1))^{-1/\gamma}$, we deduce from (3.7) and (3.10) that,

$$\|y\|_{s,t,\gamma} \leq M_2(1 + \|B\|_{0,T,\gamma}) \left(\|y\|_{s,t,\infty} + (1 + \|B\|_{0,T,\gamma}) \|y\|_{s,t,\infty} \Delta^\gamma + \|y\|_{s,t,\gamma} \Delta^\gamma \right), \quad (3.11)$$

for a large enough constant $M_2 \geq \tilde{M}(1 + |x_0|Te^{M_1T} + \|g\|_\infty + \|g'\|_\infty)$. From the estimate (3.11), a straightforward calculation (for instance, see [HN07, p. 407]) yields

$$\begin{aligned} \|y\|_{0,T,\infty} &\leq M2^{MT(1+\|B\|_{0,T,\gamma})^{1/\gamma}} |y_0|, \\ \|y\|_{0,T,\gamma} &\leq M(1 + \|B\|_{0,T,\gamma})^{2/\gamma} 2^{MT(1+\|B\|_{0,T,\gamma})^{1/\gamma}} |y_0|, \end{aligned}$$

for a large constant M , thus concluding the proof. \square

In the next sections, we will also require bounds on the *second* derivative of the solution flow with respect to its initial condition. Given an equation of the type (3.3), the second variation $z_t \in \mathbb{R}^p$ of x with respect to its initial condition satisfies an equation of the type

$$z_t = z_0 + \int_0^t h(x_r)(z_r, dB_r) + \int_0^t \hat{h}(x_r)(y_r, y_r, dB_r). \quad (3.12)$$

Here, the maps $h: \mathbb{R}^n \rightarrow \mathcal{L}_2(\mathbb{R}^p, \mathbb{R}^d; \mathbb{R}^p)$ and $\hat{h}: \mathbb{R}^n \rightarrow \mathcal{L}_3(\mathbb{R}^m, \mathbb{R}^m, \mathbb{R}^d; \mathbb{R}^p)$ are bounded with bounded derivatives. Bounds on z are covered by the following corollary to Lemma 3.2:

Corollary 3.4 *For the process z_t defined in (3.12), we have the pathwise bounds:*

$$\|z\|_{0,T,\gamma} \leq M(1 + |x_0|^5)(1 + |y_0|^2)|z_0| \exp(M\|B\|_{0,T,\gamma}^{1/\gamma})$$

where M is an absolute constant which depends only on $\|f\|_\infty, \|f'\|_\infty, \|f'_0\|_\infty, \|g\|_\infty, \|g'\|_\infty$ and T .

Proof. Denote by Z_t the $\mathbb{R}^{p \times p}$ -valued solution to the homogeneous equation

$$Z_t \xi = \xi + \int_0^t h(x_r)(Z_r \xi, dB_r), \quad \xi \in \mathbb{R}^p.$$

It follows from Lemma 3.2 that both Z and Z^{-1} satisfy the bound (3.5b), just like y did. On the other hand, z can be solved by the variation of constants formula:

$$z_t = Z_t z_0 + Z_t \int_0^t Z_s^{-1} \hat{h}(x_s)(y_s, y_s, dB_s). \quad (3.13)$$

We now use the fact that if f and g are two functions that are γ -Hölder continuous for some $\gamma > \frac{1}{2}$, then there exist constants M_T such that

$$\left\| \int_0^\cdot f(t) dg(t) \right\|_{0,T,\gamma} \leq M_T \|f\|_{0,T,\gamma} \|g\|_{0,T,\gamma} . \quad (3.14)$$

(See [You36] or for example [MN08, Eqn 2.1] for a formulation that is closer to (3.14).) The claim now follows from (3.13) by a simple application of Lemma 3.2. \square

The above two lemmas immediately give us the required *conditional* moment estimates for the Jacobian $J_{0,T}$.

Lemma 3.5 *Let $X_t, J_{0,t}, m_t$ be as defined in (3.2a), (3.2b) and (2.8) respectively. Then for any $T > 0, p \geq 1$ and $\gamma \in (1/2, H)$ we have*

$$\begin{aligned} \tilde{\mathbb{E}}\left(\|J_{0,\cdot}\|_{0,T,\gamma}^p\right) &< M e^{M(1+\|m\|_{0,T,\gamma})^{1/\gamma}} \\ \tilde{\mathbb{E}}\left(\|J_{0,\cdot}^{-1}\|_{0,T,\gamma}^p\right) &< M e^{M(1+\|m\|_{0,T,\gamma})^{1/\gamma}} \end{aligned}$$

Proof. We first write $B_t = \tilde{B}_t + m_t$ as in (2.8). Thus (3.2b) can be written as:

$$dJ_{0,t} = V'_0(X_t)J_{0,t}dt + V'(X_t)J_{0,t} \circ (d\tilde{B}_t + dm_t)$$

By applying Lemma 3.2 to the SDE in (3.2b) with $y_t = (t, \tilde{B}_t + m_t)$, $f_0 = V_0$, $f = (0, V_1, \dots, V_d)$, $g = (V'_0, \dots, V'_d)$ and using the fact that $1/\gamma > 1$ we obtain:

$$\|J\|_{0,T,\gamma}^p \leq M e^{M(1+\|\tilde{B}\|_{0,T,\gamma}^{1/\gamma} + \|m\|_{0,T,\gamma}^{1/\gamma})} . \quad (3.15)$$

The fact that $\gamma > \frac{1}{2}$ is crucial. Since \tilde{B} is a Gaussian process, it follows by Fernique's theorem [Bog98] that the Hölder norm of \tilde{B} has Gaussian tails, thus proving our first claim. An identical argument for $J_{0,t}^{-1}$ finishes the proof. \square

Remark 3.6 In an identical way, one can obtain similar bounds on the conditional moments of the solution and of its second variation with respect to its initial condition.

4 Smoothness of conditional densities

Our goal in this section is to prove that for the SDE:

$$\begin{aligned} dX_t &= V_0(X_t)dt + \sum_{i=1}^d V_i(X_t) \circ dB_t^i \\ X_0 &= x_0 \in \mathbb{R}^n \end{aligned}$$

where dB_t is a *two-sided* fBM, if the vector fields V satisfy Hörmander's celebrated Lie bracket condition (see below), then the *conditional* distribution of X_t given $\{B_s : s \leq 0\}$ almost surely has a smooth density with respect to Lebesgue measure on \mathbb{R}^n for every $t > 0$.

In [BH07], the authors studied this problem and showed that the law of X_t has a density, but with one key difference. In [BH07], the driving noise dB_t was a one-sided fBM and no conditioning was involved. As explained in (2.8), we can rewrite $B_t = \tilde{B}_t + m(t)$ and thus obtain

$$dX_t = V_0(X_t)dt + V(X_t)dm_t + V(X_t) \circ d\tilde{B}_t . \quad (4.1)$$

While the distribution of the conditioned process \tilde{B}_t is different from that of the one-sided fBm, its covariance satisfies very similar bounds so that the results from [BH07] would apply. (See Proposition 4.4 below.)

However, the difficulty comes from the derivative of the conditional mean m , which diverges at $t = 0$ like $\mathcal{O}(t^{H-1})$. To use the results from [BH07] directly, we would need the boundedness of the driving vector fields, uniformly in time, so that they are not applicable to our case. Our treatment of this singularity is inspired by [HM09], where a related situation appears due to the infinite-dimensionality of the system considered. Using an iterative argument, we will show the invertibility of a slightly modified reduced Malliavin matrix which will then, by standard arguments, yield the smoothness of laws of the SDE given in (4.1).

Before we state Hörmander's condition, let us introduce a shorthand notation. For any multiindex $I = (i_1, i_2, \dots, i_k) \in \{0, 1, 2, \dots, d\}^k$, define the commutator:

$$V_I \equiv [V_{i_1}, [V_{i_2}, \dots, [V_{i_{k-1}}, V_{i_k}], \dots]] .$$

Also define the sets of vector fields

$$\mathcal{V}_n = \left\{ V_I, I \in \{0, \dots, d\}^{n-1} \times \{1, \dots, d\} \right\}, \quad \bar{\mathcal{V}}_n = \bigcup_{k=1}^n \mathcal{V}_k . \quad (4.2)$$

With these notations at hand, we now formulate Hörmander's bracket condition [Hör67]

Assumption 4.1 *For every $x_0 \in \mathbb{R}^n$, there exists $N \in \mathbb{N}$ such that the identity*

$$\text{span}\{U(x_0), U \in \bar{\mathcal{V}}_N\} = \mathbb{R}^n . \quad (4.3)$$

holds.

4.1 Invertibility of the Malliavin Matrix

It is well-known from the works of Malliavin, Bismut, Kusuoka, Stroock and others [Mal78, Bis81, KS84, KS85, KS87, Nor86, Mal97, Nua06] that one way of proving the smoothness of laws of X_T (at least when driven by Brownian motion) is to first show the invertibility of the reduced Malliavin matrix :

$$\tilde{C}_T \stackrel{\text{def}}{=} \mathcal{A} \mathcal{A}^* = \int_0^T J_{0,s}^{-1} V(X_s) V(X_s)^* (J_{0,s}^{-1})^* ds , \quad (4.4)$$

where we defined the (random) operator: $\mathcal{A} : L^2[0, T] \mapsto \mathbb{R}^n$ by

$$\mathcal{A}v = \int_0^T J_{0,s}^{-1} V(X_s) v(s) ds . \quad (4.5)$$

However for the stochastic control argument used in Section 5 below for proving the strong Feller property, the invertibility of \tilde{C}_T turns out not to be sufficient. We thus fix now a positive smooth function $h: [0, T] \rightarrow [0, 1]$ such that

$$h(s) = \begin{cases} 1 & \text{if } s \leq T/2, \\ 0 & \text{if } s \geq 3T/4. \end{cases} \quad (4.6)$$

With this definition at hand, we introduce a modified reduced Malliavin matrix C_T given by

$$C_T \stackrel{\text{def}}{=} \mathcal{A} h^2 \mathcal{A}^* . \quad (4.7)$$

The aim of this section is to show that C_T is invertible, and to obtain moment estimates for $|C_T^{-1}|$. See Remark 5.4 below for the reason behind introducing the function h and the matrix C_T . Notice that since h is bounded, the invertibility of C_T immediately implies the invertibility of \tilde{C}_T , but the converse is of course not true in general. The main result of this section can be formulated in the following way:

Theorem 4.2 *Assume that the vector fields $\{V_i\}$ have bounded derivatives of all orders for $i \geq 0$ and are bounded for $i > 0$. Assume furthermore that Assumption 4.1 holds, fix $T > 0$ and let C_T be as in (4.7). Then, for any $R > 0$ and any $p \geq 1$, there exists a constant M such that the bound*

$$\tilde{\mathbb{P}}\left(\inf_{\|\varphi\|=1} \langle \varphi, C_T \varphi \rangle \leq \varepsilon\right) \leq M \varepsilon^p, \quad \varphi \in \mathbb{R}^n,$$

holds for all $\varepsilon \in (0, 1]$ and for all x_0 and ω such that $|x_0| + \|\omega\|_{(\gamma, \delta)} \leq R$. Here, x_0 is the initial condition of the SDE (4.1).

Corollary 4.3 *The matrix C_T is almost surely invertible and, for any $R > 0$ and any $p \geq 1$, there exists a constant M such that*

$$\tilde{\mathbb{E}}[|C_T^{-p}|] < M,$$

uniformly over all x_0 and ω such that $|x_0| + \|\omega\|_{(\gamma, \delta)} \leq R$.

Proof. This follows immediately from the fact that $\mathbb{E}K \leq \int_0^\infty \mathbb{P}(K \geq x) dx$ for any positive random variable K . \square

Before proceeding further, we need the following version of Norris' lemma [KS85, Nor86] adapted from [BH07] which will be the key technical estimate needed for showing the invertibility of C_T . In the case of white driving noise, Norris' lemma essentially states that if a semimartingale is small and if one has some *a priori* bounds on its components, then both the bounded variation part and the martingale part have to be small. In this sense, it can be viewed as a quantitative version of the Doob-Meyer semimartingale decomposition theorem. One version of this result for fractional driving noise is given by:

Proposition 4.4 *Let $H \in (\frac{1}{2}, 1)$ and $\gamma \in (\frac{1}{2}, H)$ and let \tilde{B} be as in (2.8). Then there exist exponents $q > 0$, and $\eta > 0$ such that the following statement holds.*

Consider any two families $a^\varepsilon, b^\varepsilon$ of processes taking values in \mathbb{R} and \mathbb{R}^d respectively such that $\mathbb{E}(\|a^\varepsilon\|_{0,T,\gamma}^p + \|b^\varepsilon\|_{0,T,\gamma}^p) < M_p \varepsilon^{-\eta p}$, uniformly in $\varepsilon \in (0, 1]$, for some $M_p > 0$ and for every $p \geq 1$. Let y^ε be the process defined by

$$y_t^\varepsilon = \int_0^t a_s^\varepsilon ds + \int_0^t b_s^\varepsilon \circ d\tilde{B}_s, \quad t \in [0, T].$$

Then, the estimate

$$\mathbb{P}\left(\|y\|_{0,T,\infty} < \varepsilon \text{ and } \|a\|_{0,T,\infty} + \|b\|_{0,T,\infty} > \varepsilon^q\right) < C_p \varepsilon^p$$

holds for every $p > 0$, uniformly for $\varepsilon \in (0, 1]$.

Proof. Setting $f(s, t) = \mathbb{E}|\tilde{B}(t) - \tilde{B}(s)|^2$ as in [BH07], we have

$$f(s, t) = \frac{t^{2H}}{2H-1} + \frac{s^{2H}}{2H-1} - 2 \int_0^s (t-r)^{H-\frac{1}{2}}(s-r)^{H-\frac{1}{2}} dr .$$

(We use the convention $s < t$.) A somewhat lengthy but straightforward calculation shows that this can be rewritten as

$$f(s, t) = |t-s|^{2H} \left(\frac{1}{2H-1} + \int_0^{\frac{s \wedge t}{|t-s|}} ((1+x)^{H-\frac{1}{2}} - x^{H-\frac{1}{2}})^2 dx \right) .$$

Similarly, one has

$$\partial_s \partial_t f(s, t) = -2 \left(H - \frac{1}{2} \right)^2 |t-s|^{2H-2} \int_0^{\frac{s \wedge t}{|t-s|}} (1+x)^{H-\frac{3}{2}} x^{H-\frac{3}{2}} dx .$$

Since the integrands of both of the integrals appearing in these expressions decay like x^{2H-3} , they are integrable so that \tilde{B} is indeed a Gaussian process of ‘type H ’, see [BH07, Eqn 3.5].

The proof is now a slight modification of that of [BH07, Prop 3.4]. It follows from [BH07, Equ 3.18] and the last displayed equation in the proof of [BH07, Prop 3.4] that there exist exponents q and β such that the bound

$$\mathbb{P} \left(\|y\|_{0,T,\infty} < \varepsilon \text{ and } \|a\|_{0,T,\infty} + \|b\|_{0,T,\infty} > \varepsilon^q \right) \leq C_p \varepsilon^p + \mathbb{P}(\|b\|_\gamma^2 + \|a\|_\gamma^2 > \varepsilon^{-\beta})$$

holds. The claim now follows from a simple application of Chebychev’s inequality, provided that we ensure that η is such that $2\eta < \beta$. \square

For making further calculations easier, we now introduce some definitions that are inspired by the proofs in [HM09].

Definition 4.5 A family of events $A \equiv \{A_\varepsilon\}_{\varepsilon \leq 1}$ of a probability space is called a “family of negligible events” if, for every $p \geq 1$ there exists a constant C_p such that $\mathbb{P}(A_\varepsilon) \leq C_p \varepsilon^p$ for every $\varepsilon \leq 1$.

Definition 4.6 Given a statement Φ_ε , we will say that “ Φ_ε holds modulo negligible events”, if there exists a family $\{A_\varepsilon\}$ of negligible events such that, for every $\varepsilon \leq 1$, the statement Φ_ε holds on the complement of A_ε .

With the above definitions, Theorem 4.2 can be restated as:

Theorem 4.7 Fix $R > 0$. Then, modulo negligible events, the bound

$$\inf_{\|\varphi\|=1} \langle \varphi, C_T \varphi \rangle \geq \varepsilon$$

holds for all x_0 and ω such that $|x_0| + \|\omega\|_{(\gamma,\delta)} \leq R$.

Remark 4.8 The family of events in question can (and will in general) depend on R , but once R is fixed the bounds are uniform over $|x_0| + \|\omega\|_{(\gamma,\delta)} \leq R$. While this uniformity is not really required in this section, it will be crucial in Section 5.

Proof of Theorem 4.7. It suffices to show that for any fixed φ with $\|\varphi\| = 1$, the bound $\langle \varphi, C_T \varphi \rangle \geq \varepsilon$ holds modulo a family of negligible events. The conclusion then follows by a standard covering argument, see for example [Nor86, p. 127]. Furthermore, throughout all of this proof, we will denote by M a constant that can change from one expression to the next and only depends on R and on the vector fields V defining the problem.

Fix now such a φ and assume “by contradiction” that $\langle \varphi, C_T \varphi \rangle < \varepsilon$. We will show that this assumption leads to the conclusion that

$$\sup_{U \in \bar{\mathcal{V}}_N} |\langle \varphi, U(x_0) \rangle| < \varepsilon^{q'}, \quad \text{for some } q' > 0,$$

modulo negligible events. However, if Hörmander’s condition is satisfied, the conclusion above cannot hold since $\{U(x_0), U \in \bar{\mathcal{V}}_N\}$ span \mathbb{R}^n so that there exists a constant c such that $\sup_{U \in \bar{\mathcal{V}}_N} |\langle \varphi, U(x_0) \rangle| \geq c$ almost surely, thus leading to a contradiction and concluding the proof. For any smooth vector field U , let us introduce the process $Z_U(t) = J_{0,t}^{-1} U(X_t)$. With this notation, we have

Lemma 4.9 *There exists $r > 0$ such that, modulo negligible events, the implication*

$$\langle \varphi, C_T \varphi \rangle < \varepsilon \quad \Rightarrow \quad \|\langle \varphi, Z_U(\cdot) \rangle\|_{0,T/2,\infty} < \varepsilon^r$$

holds for all $U \in \mathcal{V}_1$.

Proof. By the definition of h , we have the bound

$$\langle \varphi, C_T \varphi \rangle = \sum_{i=1}^d \int_0^T \langle \varphi, J_{0,s}^{-1} V_i(X_s) h(s) \rangle^2 ds \geq \sum_{i=1}^d \int_0^{T/2} \langle \varphi, J_{0,s}^{-1} V_i(X_s) \rangle^2 ds. \quad (4.8)$$

We now wish to turn this into a lower bound of order ε^r (for some $r > 0$) on the supremum of the integrand. Our main tool for this is the following interpolation lemma, the proof of which is postponed to the appendix:

Lemma 4.10 *There exists $M > 0$ such that*

$$\|f\|_{0,T,\infty} \leq M \max \left(\|f\|_{L^2[0,T]}, \|f\|_{L^2[0,T]}^{\frac{2\gamma}{2\gamma+1}} \|f\|_{0,T,\gamma}^{\frac{1}{2\gamma+1}} \right),$$

holds for every γ -Hölder continuous function $f : [0, T] \mapsto \mathbb{R}$.

By Lemma 3.5 and Chebyshev’s inequality it follows that $\|\langle \varphi, J_{0,\cdot}^{-1} V(X_\cdot) \rangle\|_{0,T,\gamma} < \varepsilon^{-1/2}$ modulo negligible events. Thus if $\langle \varphi, C_T \varphi \rangle < \varepsilon$, from Lemma 4.10 and (4.8) we deduce that, for each $1 \leq j \leq d$

$$\|\langle \varphi, J_{0,\cdot}^{-1} V_j(X_\cdot) \rangle\|_{0,T/2,\infty} < M \varepsilon^{\frac{4\gamma-1}{2(2\gamma+1)}}, \quad (4.9)$$

modulo negligible events. The claim follows by choosing $r < \frac{4\gamma-1}{2(2\gamma+1)}$. \square

We will now use an induction argument to show that a similar bound holds with V_j replaced by any vector field obtained by taking finitely many Lie brackets between the V_j ’s. Whilst the gist of the argument is standard, see [KS85, Nor86, BH07], a technical difficulty arises from the divergence of the derivative of $m(t)$ at $t = 0$. To surmount this

difficulty, we take inspiration from [HM09] and introduce a (small) parameter $\alpha > 0$ to be specified at the end of the iteration. Then, by Lemma 2.5, we have

$$\left\| \frac{dm_\cdot}{dt} \right\|_{\varepsilon^\alpha, T/2, \infty} \leq M \varepsilon^{\alpha(H-1)}. \quad (4.10)$$

Instead of considering supremums over the time interval $[0, T/2]$, we will now consider supremums over $[\varepsilon^\alpha, T/2]$ instead, thus avoiding the singularity.

In order to formalise this, for $r > 0$ and $i \geq 1$, define the events

$$K_i^\varepsilon(\alpha, r) = \left\{ \|\langle \varphi, Z_U(\cdot) \rangle\|_{\varepsilon^\alpha, T/2, \infty} < \varepsilon^r, \quad \forall U \in \mathcal{V}_i \right\}.$$

With this notation at hand, the recursion step in our argument can be formulated as follows:

Lemma 4.11 *There exists $\bar{q} > 0$ such that, for every $\alpha > 0$ and every $i \geq 1$, the implication*

$$K_i^\varepsilon(\alpha, r) \Rightarrow K_{i+1}^\varepsilon(\alpha, \bar{q}r + \alpha(H-1)),$$

holds modulo negligible events.

Proof. Assume that $K_i^\varepsilon(\alpha, r)$ holds. For any $t \leq T/2$, it then follows from the chain rule that

$$\begin{aligned} \langle \varphi, Z_U(t) - Z_U(\varepsilon^\alpha) \rangle &= \int_{\varepsilon^\alpha}^t \langle \varphi, J_{0,s}^{-1}[V_0, U](X_s) \rangle ds \\ &+ \sum_{i=1}^d \int_{\varepsilon^\alpha}^t \langle \varphi, J_{0,s}^{-1}[V_i, U](X_s) \rangle dm_s^i + \sum_{i=1}^d \int_{\varepsilon^\alpha}^t \langle \varphi, J_{0,s}^{-1}[V_i, U](X_s) \rangle \circ d\tilde{B}_s^i, \end{aligned} \quad (4.11)$$

where \tilde{B}_\cdot is as in (2.8).

Now by the hypothesis, Proposition 4.4 and (4.11), it follows that for any $U \in \mathcal{V}_i$,

$$\left\| \left\langle \varphi, \sum_{i=1}^d J_{0,\cdot}^{-1}[V_i, U](X_\cdot) \right\rangle \right\|_{\varepsilon^\alpha, T/2, \infty} < M \varepsilon^{rq}, \quad (4.12a)$$

$$\left\| \left\langle \varphi, J_{0,\cdot}^{-1}[V_0, U](X_\cdot) + \sum_{i=1}^d \frac{dm_\cdot^i}{ds} J_{0,\cdot}^{-1}[V_i, U](X_\cdot) \right\rangle \right\|_{\varepsilon^\alpha, T/2, \infty} < M \varepsilon^{rq}, \quad (4.12b)$$

modulo negligible events, with q as given by Proposition 4.4. Plugging (4.10) and (4.12a) into (4.12b), it follows that for every $U \in \mathcal{V}_i$,

$$\|\langle \varphi, J_{0,\cdot}^{-1}[V_0, U](X_\cdot) \rangle\|_{\varepsilon^\alpha, T/2, \infty} < M \varepsilon^{rq + \alpha(H-1)}$$

modulo negligible events. Choosing any $\bar{q} < q$, the claim then follows from the fact that for every exponent $\delta > 0$, one has $M \varepsilon^\delta < 1$ modulo negligible events. \square

Now we iterate the previous argument. Let $N \in \mathbb{N}$ be such that

$$\inf_{|x| \leq R} \inf_{|\xi| = 1} \sum_{U \in \mathcal{V}_N} |\langle \xi, U(x) \rangle|^2 > 0. \quad (4.13)$$

It follows from the Hörmander condition and the fact that the vector fields are smooth that such an N exists. From Lemmas 4.9 & 4.11, we obtain that for every $U \in \bar{\mathcal{V}}_N$, modulo negligible events,

$$\|\langle \varphi, J_{0,\cdot}^{-1} U(X_\cdot) \rangle\|_{\varepsilon^\alpha, T/2, \infty} < \varepsilon^{q_1}, \quad (4.14)$$

where $q_1 = r\bar{q}^N + \alpha(H-1)\frac{(1-r\bar{q})^N}{1-r\bar{q}}$. Since \bar{q} and r are fixed, we can choose α small enough such that $q_1 > 0$.

Now we make use of the Hölder continuity of $J_{0,\cdot}^{-1}$ and X . Indeed, it follows from Lemma 3.5 that

$$\|J_{0,\cdot}^{-1} U(X_\cdot) - U(x_0)\|_{0, \varepsilon^\alpha, \infty} < \varepsilon^{q_2}, \quad (4.15)$$

modulo negligible events, for any $q_2 < \alpha\gamma$. Combining (4.14) and (4.15), we conclude that

$$\sup_{U \in \bar{\mathcal{V}}_N} |\langle \varphi, U(x_0) \rangle| < \varepsilon^{q_1} + \varepsilon^{q_2},$$

modulo negligible events. This is in direct contradiction with (4.13), thus concluding the proof of Theorem 4.7. \square

We conclude that the Malliavin matrix \tilde{C}_T defined in (4.4) is almost surely invertible and $\tilde{\mathbb{E}}(|\tilde{C}_T|^{-p}) < M$ for $p \geq 1$. As a consequence of Theorem 4.2, we obtain the smoothness of the laws of X_t conditional on a given past $\omega \in \mathcal{W}$:

Theorem 4.12 *Let $\varphi_t(x, \tilde{\omega})$ be the solution to (1.1) with notations as in Section 2.1, and assume that Assumption 4.1 holds. Then, for every $\omega \in \mathcal{W}$, the distribution of $\varphi_t(x, \tilde{\omega})$ under $\mathcal{P}(\omega, \cdot)$ has a smooth density with respect to Lebesgue measure.*

Proof. In order to obtain the smoothness of the densities, it suffices by [Nua06, Mal97, BH07, NS09] to show that the Malliavin matrix \mathcal{M}_T of the map $W \mapsto X_T^x$, where W is the standard Wiener process describing the solution as in (2.8), is invertible with negative moments of all orders. As we will see in Section 5 below (see equation 5.3), \mathcal{M}_T is given by

$$\langle \xi, \mathcal{M}_T \xi \rangle = \|(\mathcal{I}^{H-\frac{1}{2}})^* \mathcal{A}^* J_{0,T}^* \xi\|_{L^2[0,T]}^2,$$

where the L^2 -adjoint of $\mathcal{I}^{H-\frac{1}{2}}$ is given by

$$((\mathcal{I}^{H-\frac{1}{2}})^* g)(t) = \int_t^T (s-t)^{H-\frac{3}{2}} g(s) ds.$$

We already know from Theorem 4.2 that $\|\mathcal{A}^* J_{0,T}^* \xi\|_{L^2[0,T]}^2 \geq \varepsilon$ modulo negligible events and we would like to make use of this information in order to obtain a lower bound on \mathcal{M}_T .

Since $\mathcal{I}^{H-\frac{1}{2}}$ and $\mathcal{D}^{H-\frac{1}{2}}$ are inverses of each other, for any measurable f we have,

$$\begin{aligned} \|f\|_{L^2[0,T]}^2 &= \langle f, \mathcal{I}^{H-\frac{1}{2}} \mathcal{D}^{H-\frac{1}{2}} f \rangle_{L^2[0,T]} \leq \|(\mathcal{I}^{H-\frac{1}{2}})^* f\|_{L^2[0,T]} \|\mathcal{D}^{H-\frac{1}{2}} f\|_{L^2[0,T]} \\ &\leq M \|(\mathcal{I}^{H-\frac{1}{2}})^* f\|_{L^2[0,T]} \|f\|_{0,T,\gamma}. \end{aligned}$$

Thus we obtain,

$$\|(\mathcal{I}^{H-\frac{1}{2}})^* \mathcal{A}^* J_{0,T}^* \xi\|_{L^2[0,T]} \geq M \|\mathcal{A}^* J_{0,T}^* \xi\|_{L^2[0,T]}^2 \|\mathcal{A}^* J_{0,T}^* \xi\|_{0,T,\gamma}^{-1}.$$

Since we know from the results in Section 3 that $\|\mathcal{A}^* J_{0,T} \xi\|_{0,T,\gamma} \leq \varepsilon^{-1}$ modulo negligible events, we conclude that

$$\|(\mathcal{I}^{H-\frac{1}{2}})^* \mathcal{A}^* J_{0,T}^* \xi\|_{L^2[0,T]} \geq \|\mathcal{A}^* J_{0,T}^* \xi\|_{L^2[0,T]}^2 \varepsilon ,$$

holds modulo negligible events. The proof now follows from Theorem 4.2. \square

5 Strong Feller Property

Recall from the arguments in Section 2.2 that the strong Feller property (2.2) is the only ingredient required for the proof of Theorem 1.3. We will actually show something slightly stronger than just (2.2). Fix some arbitrary final time $T > 0$, a Fréchet differentiable bounded map $\psi: \mathcal{C}([1, T], \mathbb{R}^n) \rightarrow \mathbb{R}$ with bounded derivative, denote by $R_{1,T}: \mathcal{C}(\mathbb{R}_+, \mathbb{R}^n) \rightarrow \mathcal{C}([1, T], \mathbb{R}^n)$ the restriction operator, and set as before

$$\bar{\mathcal{Q}}\psi(x, \omega) \stackrel{\text{def}}{=} \int_{\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)} \psi(R_{1,T} z) \bar{\mathcal{Q}}(x, \omega; dz) .$$

With these notations, the main result of this article is:

Theorem 5.1 *Assume that the vector fields $\{V_i\}_{i \geq 0}$ have bounded derivatives of all orders and that furthermore the vector fields $\{V_i\}_{i \geq 1}$ are bounded. Then, if Hörmander's bracket condition (Assumption 4.1) holds, there exists a continuous function $K: \mathbb{R}^n \rightarrow \mathbb{R}_+$ and a constant $M > 0$ such that the bound*

$$|D_x \bar{\mathcal{Q}}\psi(x, \omega)| \leq K(x) e^{M \|\mathcal{G}\omega\|_{0,T,\gamma}^{1/\gamma}} \|\psi\|_{\infty} , \quad (5.1)$$

holds uniformly over all test functions ψ as above. Here, \mathcal{G} was defined in (2.6) and both K and M are independent of T and of ψ . In particular, (1.1) is strong Feller.

The remainder of this section is devoted to the proof of this result. For conciseness, from now onwards we will use the shorthand notation

$$|\Psi(x, \omega)| \leq E(x, \omega) ,$$

to indicate that there exist a continuous function K on \mathbb{R}^n and a constant M such that the expression $\Psi(x, \omega)$ is bounded by the right hand side of (5.1).

To obtain the bound for $D \bar{\mathcal{Q}}\psi(\cdot, \omega)$, we use a ‘stochastic control’ argument which is inspired by [HM06, HM09] and is similar in spirit to the arguments based on the Bismut-Elworthy-Li formula [EL94, DPEZ95]. Let $X_t, J_{0,t}$ be the solution and its Jacobian, as defined in equations (3.2a) and (3.2b) respectively, and let $\varphi_t(x, \tilde{\omega}) = X_t^x$ denote the flow map as in Section 2.1. By the variation of constants formula, the Fréchet derivative of the flow map in the direction of $\int_0^\cdot v(s) ds$ is then given by:

$$D_v X_t^x = D_v \varphi_t(x, \tilde{\omega}) = J_{0,t} \mathcal{A} v \quad (5.2)$$

Remark 5.2 Notice that the operator D defined above is a standard Fréchet derivative unlike in the case of $H = \frac{1}{2}$ where it is the Malliavin derivative. This is because the integration with respect to the fBM for $H > \frac{1}{2}$ is just the Reimann-Stieltjes integral, whereas for $H = \frac{1}{2}$ the stochastic integral is not even a continuous (let alone differentiable) map of the noise in the case of multiplicative noise.

As already mentioned in Remark 2.4, the operator $\mathcal{D}^{H-\frac{1}{2}}$ and its inverse $\mathcal{I}^{H-\frac{1}{2}}$ provides an isometry between the Cameron-Martin space of the Gaussian measure $\mathcal{P}(0, \cdot)$ and that of Wiener measure.

As a consequence, if v is such that $\tilde{v} = \mathcal{D}^{H-1/2}v \in L^2(\mathbb{R}_+, \mathbb{R}^d)$, then we have the identity

$$\mathbf{D}_{\tilde{v}} X_t^x = D_v \varphi_t(x, \tilde{\omega}) = J_{0,t} \mathcal{A}v = J_{0,t} \mathcal{A} \mathcal{I}^{H-\frac{1}{2}} \tilde{v}, \quad (5.3)$$

where \mathbf{D} is the Malliavin derivative of X_t^x when interpreted as a function on Wiener space via the representation (2.8).

Remark 5.3 It follows from (5.3) and the definition of $\mathcal{I}^{H-\frac{1}{2}}$ that for some constant c_H one has the identity

$$\mathbf{D}_s X_t^x = c_H \int_s^t J_{r,s} V(X_r^x) (r-t)^{H-3/2} dr, \quad (5.4)$$

for $s \leq t$. Here, V denotes the matrix (V_1, \dots, V_d) . Note also that $\mathbf{D}_s X_t^x = 0$ for $s > t$ since X is an adapted process.

5.1 Stochastic control

The central idea of the stochastic control argument is as follows. Given an initial displacement $\xi \in \mathbb{R}^n$, we seek for a ‘control’ v on the time interval $[0, 1]$ that solves $\mathcal{A}v = \xi$. If this can be achieved, then we extend v to all of \mathbb{R}_+ by setting $v(s) = 0$ for $s \geq 1$ and we define $\tilde{v} = \mathcal{D}^{H-1/2}v$. If v is sufficiently regular in time so that this makes sense, then it follows that we have the identity

$$\mathbf{D}_{\tilde{v}} X_T^x = J_{0,T} \xi, \quad (5.5)$$

for every $T \geq 1$. In particular, for a \mathcal{C}^1 test function ψ as above and for $\xi \in \mathbb{R}^n$,

$$\begin{aligned} \langle D\bar{Q}\psi(x, \omega), \xi \rangle &= \tilde{\mathbb{E}} \left((D\psi)(\varphi(x, \tilde{\omega})) J_{0,\cdot} \xi \right) \\ &= \tilde{\mathbb{E}} \left((D\psi)(\varphi(x, \tilde{\omega})) \mathbf{D}_{\tilde{v}} X \right) = \tilde{\mathbb{E}} \left(\mathbf{D}_{\tilde{v}} (\psi(\varphi(x, \tilde{\omega}))) \right). \end{aligned} \quad (5.6a)$$

It then follows from the integration by parts formula of Malliavin calculus [Nua06] that

$$\begin{aligned} \dots &= \tilde{\mathbb{E}} \left(\psi(\varphi_T(x, \tilde{\omega})) \int_0^T \tilde{v}(s) dW_s \right) = \tilde{\mathbb{E}} \left(\psi(\varphi_T(x, \tilde{\omega})) \int_0^T \mathcal{D}^{H-1/2} v(s) dW_s \right) \\ &\leq \|\psi\|_\infty \tilde{\mathbb{E}} \left| \int_0^T \mathcal{D}^{H-1/2} v(s) dW_s \right|, \end{aligned} \quad (5.6b)$$

which is our main bound on the way to (5.1). Notice that the stochastic integral appearing in the expression above is a Skorohod integral, since the integrand is in general not adapted to the filtration generated by W . In order to have the identity (5.1), it therefore remains to find a control v satisfying $\mathcal{A}v = \xi$ and such that

$$\tilde{\mathbb{E}} \left(\int_0^\infty (\mathcal{D}^{H-\frac{1}{2}} v)(s) dW_s \right)^2 < E(x, \omega). \quad (5.7)$$

A natural candidate for such a control v can be identified by a simple least squares formula. Indeed, note that the adjoint $\mathcal{A}^* : \mathbb{R}^n \mapsto L^2([0, \infty), \mathbb{R}^n)$ of \mathcal{A} is given by:

$$(\mathcal{A}^*\xi)(s) = V(X_s)^*(J_{0,s}^{-1})^*\xi, \quad \xi \in \mathbb{R}^n. \quad (5.8)$$

Assuming that $\mathcal{A}h\mathcal{A}^*$ is invertible (recall that $h: [0, 1] \rightarrow \mathbb{R}_+$ was defined in (4.6)), a possible solution to the equation $\mathcal{A}v = \xi$ is given by

$$v(s) \equiv h(s) (\mathcal{A}^*(\mathcal{A}h\mathcal{A}^*)^{-1}\xi)(s). \quad (5.9)$$

Note that $\mathcal{A}h\mathcal{A}^*$ is precisely equal to the modified Malliavin matrix defined in (4.7), so that this expression does indeed make sense by the results from the previous section.

Remark 5.4 For the identity (5.3) to hold, the stochastic control $v(\cdot)$ needs to belong to the Cameron-Martin space of $\mathcal{P}(0, \cdot)$, and the function $h(\cdot)$ is introduced just for this purpose. Indeed, since we extend the control $v(\cdot)$ to vanish outside the interval $[0, 1]$, one needs to be careful about the regularity of $v(s)$ at $s = 1$. It can be shown using the fractional integrals that, in order for $v(\cdot)$ to have the required regularity, the function $h(s)$ should be smaller than $\mathcal{O}((1-s)^H)$ for $s \approx 1$. The specific $h(\cdot)$ we chose in (4.6), of course, satisfies this.

In the rest of this section we show that (5.7) holds for the the control v defined in (5.9) and thus establish the strong Feller property of the SDE defined in (4.1). We start with the following estimate which will be used repeatedly.

Lemma 5.5 *The stochastic control v defined in (5.9) is continuous and satisfies the bound*

$$\tilde{\mathbb{E}}\|v\|_{0,1,\gamma} < \mathbb{E}(x, \omega).$$

Proof. Setting $v(s) = h(s)u(s)$, we obtain from (5.9) the bound

$$\|u\|_{0,1,\gamma} \leq \|\mathcal{A}^*(\mathcal{A}h\mathcal{A}^*)^{-1}\xi\|_{0,1,\gamma}. \quad (5.10)$$

From (5.8), Lemma 3.5 and the fact that $V \in C^\infty$ it follows that

$$\|\mathcal{A}^*\eta\|_{0,1,\gamma} \leq M\|V(X_\cdot)\|_{0,1,\gamma}\|J_{0,\cdot}^{-1}\|_{0,1,\gamma}|\eta|.$$

Combining this with (5.10) and Cauchy-Schwarz, we deduce that

$$\tilde{\mathbb{E}}\|u\|_{0,1,\gamma} \leq \mathbb{E}(x, \omega)\tilde{\mathbb{E}}|(\mathcal{A}h\mathcal{A}^*)^{-1}\xi|^2 \leq \mathbb{E}(x, \omega),$$

where the last estimate follows from Theorem 4.2. Since h vanishes in a neighbourhood of $t = 1$, the claim follows. \square

Remark 5.6 If we had not introduced the control h , we could still obtain continuity of v on $[0, 1]$, but we would have a discontinuity at $t = 1$.

As a consequence, we show that $\mathcal{D}^{H-1/2}v$ is square integrable:

Theorem 5.7 *For the stochastic control $v(\cdot)$ defined in (5.9) and the operator \mathcal{D} defined in (2.4), the following estimate holds:*

$$\tilde{\mathbb{E}}\left(\int_0^\infty |\mathcal{D}^{H-\frac{1}{2}}v(s)|^2 ds\right) < \mathbb{E}(x, \omega).$$

Remark 5.8 One might think that Theorem 5.7 is sufficient to show (5.7). This is not the case since \tilde{v} is not adapted to the filtration generated by W and so Itô's isometry does not hold.

Before we turn to the proof of Theorem 5.7, we state the following useful result:

Lemma 5.9 For any $s \geq 0$, $\alpha > 0$ and $f \in C^\gamma$ with $\gamma > \alpha$,

$$\begin{aligned} \Gamma(1-\alpha) \mathcal{D}^\alpha f(s) &= \frac{d}{ds} \int_0^s (s-r)^{-\alpha} f(r) dr \\ &= s^{-\alpha} f(s) - \alpha \int_0^s (s-r)^{-\alpha-1} (f(r) - f(s)) dr. \end{aligned} \quad (5.11)$$

Proof. See [SKM93, Thm 13.1], as well as Lemma 13.1 in the same monograph. \square

Remark 5.10 The right hand side of equation (5.11) is called the Marchaud fractional derivative. Lemma 5.9 actually holds true for a larger class of functions f , namely those such that $f = \mathcal{I}^\alpha g$ for some $g \in L^1([0, 1], \mathbb{R})$.

Proof of Theorem 5.7. The main idea behind this proof, and a recurrent theme in this section is the following. Although the stochastic control $v(\cdot)$ vanishes outside $[0, 1]$, the integrand $\mathcal{D}^{H-\frac{1}{2}} v(s)$ is non-zero for all $s \in [0, \infty)$. Thus the square integrability of $\mathcal{D}^{H-\frac{1}{2}} v(s)$ on the interval $[0, \infty)$ is shown by dealing with the singularities of the integrand differently at $s = 0$, and $s = \infty$. At $s = 0$, we use the fact that v is Hölder continuous whereas at $s = \infty$, we use the fact that v is bounded.

By definition,

$$\mathcal{D}^{H-\frac{1}{2}} v(s) = \frac{d}{ds} \int_0^s (s-r)^{\frac{1}{2}-H} v(r) dr.$$

Applying Lemma 5.9 with $\alpha = H - \frac{1}{2}$, we obtain the bound

$$|\mathcal{D}^{H-\frac{1}{2}} v(s)| \leq M s^{\frac{1}{2}-H} |v(s)| + M \int_0^s (s-r)^{-\frac{1}{2}-H} |v(r) - v(s)| dr,$$

for some constant $M > 0$. Noting that the control $v(\cdot)$ vanishes outside the interval $[0, 1]$, we have

$$\begin{aligned} \tilde{\mathbb{E}} \left(\int_0^\infty |\mathcal{D}^{H-\frac{1}{2}} v(s)|^2 ds \right) &\leq M \tilde{\mathbb{E}} \|v\|_{0,1,\infty}^2 \int_0^1 s^{1-2H} ds \\ &\quad + M \tilde{\mathbb{E}} \int_0^\infty \left(\int_0^s (s-r)^{-\frac{1}{2}-H} |v(r) - v(s)| dr \right)^2 ds. \end{aligned} \quad (5.12)$$

The first integral $\int_0^1 s^{1-2H} ds$ in equation (5.12) is clearly finite since $H < 1$, so that this term is bounded by $\mathbb{E}(x, \omega)$ by Lemma 5.5. Now we show that the second term has a similar bound. To this end, set

$$\zeta(s) \equiv \int_0^s (s-r)^{-\frac{1}{2}-H} |v(r) - v(s)| dr.$$

For s smaller than 1, we obtain

$$|\zeta(s)| = \int_0^s (s-r)^{-\frac{1}{2}-H} |v(r) - v(s)| dr \leq M \|v\|_{0,1,\gamma} \int_0^s s^{-\frac{1}{2}-H+\gamma} ds$$

$$\leq M \|v\|_{0,1,\gamma} s^{\frac{1}{2}-H+\gamma},$$

so that the required bound follows again from Lemma 5.5. For $s \geq 1$, we observe that $\zeta(s)$ has a square integrable singularity, and thus the boundedness of the control $v(\cdot)$ will be sufficient. Indeed it follows from the definition of h (4.6) that

$$|\zeta(s)| \leq \int_0^{3/4} (s-3/4)^{-\frac{1}{2}-H} |v(r) - v(s)| dr \leq M \|v\|_{0,1,\infty} (s-3/4)^{-\frac{1}{2}-H}.$$

Since this expression is square integrable, it follows that $\tilde{\mathbb{E}} \int_0^\infty |\zeta^2(s)| ds < E(x, \omega)$, and the claim follows from (5.12). \square

The next theorem is more involved and we prove it via a series of lemmas.

Theorem 5.11 *Let v be as defined in (5.9), the operator \mathcal{D} as defined in (2.4) and \mathbf{D} be the Malliavin derivative. Then following estimate holds:*

$$\tilde{\mathbb{E}} \left(\int_0^\infty \int_0^\infty \mathbf{D}_t(\mathcal{D}^{H-\frac{1}{2}}v(s)) \mathbf{D}_s(\mathcal{D}^{H-\frac{1}{2}}v(t)) ds dt \right) < E(x, \omega). \quad (5.13)$$

Proof. The main difficulty here is to obtain control over the modulus of continuity of the quantity $\mathbf{D}_s v(t)$ for all $s, t \in [0, 1]$. On route to establishing the Hölder continuity of the Malliavin derivative applied to the stochastic control, we need to perform several careful estimates using the fractional derivatives.

First notice that, since $v(s) \in \mathcal{F}_1$ ($\mathcal{F}_t, t \geq 0$ denotes the filtration generated by the increments of W), we have $\mathbf{D}_t(\mathcal{D}^{H-\frac{1}{2}}v(s)) = 0$ for $t > 1$, so that the two integrals in (5.13) stop at 1. We thus have

$$\begin{aligned} \tilde{\mathbb{E}} \left(\int_0^\infty \int_0^\infty \mathbf{D}_t(\mathcal{D}^{H-\frac{1}{2}}v(s)) \mathbf{D}_s(\mathcal{D}^{H-\frac{1}{2}}v(t)) ds dt \right) \\ = \tilde{\mathbb{E}} \left(\int_0^1 \int_0^1 (\mathcal{D}^{H-\frac{1}{2}}\mathbf{D}_t v)(s) (\mathcal{D}^{H-\frac{1}{2}}\mathbf{D}_s v)(t) ds dt \right) \\ \leq \tilde{\mathbb{E}} \left(\int_0^1 \int_0^1 |(\mathcal{D}^{H-\frac{1}{2}}\mathbf{D}_t v)(s)|^2 dt ds \right). \end{aligned}$$

By Lemma 5.9, the claim follows if we can show that to show

$$\tilde{\mathbb{E}} \left(\int_0^1 \int_0^1 \left| \int_0^s (s-r)^{-\frac{1}{2}-H} (\mathbf{D}_t v(s) - \mathbf{D}_t v(r)) dr \right|^2 dt ds \right) < E(x, \omega). \quad (5.14)$$

Expanding the integrand, we obtain

$$\begin{aligned} \left| \mathbf{D}_t \left(\int_0^s (s-r)^{-\frac{1}{2}-H} (v(s) - v(r)) dr \right) \right| &\leq \int_0^t (s-r)^{-\frac{1}{2}-H} |\mathbf{D}_t v(s)| dr \\ &\quad + \left| \mathbf{D}_t \left(\int_t^s (s-r)^{-\frac{1}{2}-H} (v(s) - v(r)) dr \right) \right| \\ &\leq |s^{\frac{1}{2}-H} - (s-t)^{\frac{1}{2}-H}| \|\mathbf{D}_t v(s)\|_{0,1,\infty} \\ &\quad + \left| \int_t^s (s-r)^{-\frac{1}{2}-H} \mathbf{D}_t(v(s) - v(r)) dr \right|. \end{aligned} \quad (5.15)$$

Since the double integral of $|s^{\frac{1}{2}-H} - (s-t)^{\frac{1}{2}-H}|^2$ is finite, the first term will be bounded by $E(x, \omega)$ if we can show that $\tilde{\mathbb{E}} \sup_{t \in [0,1]} \|\mathbf{D}_t v\|_{0,1,\infty}^2 < E(x, \omega)$. Since this

bound will be a byproduct of the analysis of the second term, we postpone it to (5.21) below.

The second term of equation (5.15) needs more delicate calculations. We will show that there exists a random variable K and an exponent $\alpha < 1 + \gamma - H$ such that the following bounds hold:

$$|\mathbf{D}_t v(s) - \mathbf{D}_t v(r)| \leq K |s - r|^\gamma |r - t|^{-\alpha}, \quad \text{with } \tilde{\mathbb{E}} K^2 < \mathbb{E}(x, \omega). \quad (5.16)$$

Assuming these bounds, we proceed similarly to the proof of Theorem 5.7:

$$\begin{aligned} & \tilde{\mathbb{E}} \left(\int_0^1 \int_0^1 \left| \int_t^s (s-r)^{-\frac{1}{2}-H} \mathbf{D}_t(v(s) - v(r)) dr \right|^2 dt ds \right) \\ & \leq \tilde{\mathbb{E}} \left(K^2 \int_0^1 \int_0^1 \left| \int_t^s (s-r)^{-\frac{1}{2}-H+\gamma} (r-t)^{-\alpha} dr \right|^2 dt ds \right) \\ & \leq \tilde{\mathbb{E}} \left(K^2 \int_0^1 \int_0^1 (s-t)^{1+2\gamma-2H-2\alpha} dt ds \right) < \mathbb{E}(x, \omega), \end{aligned}$$

where we used the bound on α to obtain the last inequality. It therefore remains to show (5.16).

The Malliavin derivative of v is given by

$$\mathbf{D}_t v(s) = h(s) \mathbf{D}_t \left(V(X_s)^* (J_{0,s}^{-1})^* (AhA^*)^{-1} \xi \right).$$

Applying the product rule, we break this into a sum of three terms:

$$\begin{aligned} \mathbf{D}_t v(s) &= h(s) \left[\mathbf{D}_t \left(V(X_s)^* \right) (J_{0,s}^{-1})^* (AhA^*)^{-1} \xi + V(X_s)^* \mathbf{D}_t \left((J_{0,s}^{-1})^* \right) (AhA^*)^{-1} \xi \right. \\ &\quad \left. + V(X_s)^* (J_{0,s}^{-1})^* \mathbf{D}_t \left((AhA^*)^{-1} \xi \right) \right] \\ &\stackrel{\text{def}}{=} \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3. \end{aligned} \quad (5.17)$$

To make the proof compact, we introduce the stochastic process Y_s given by collecting the various objects appearing in this expression:

$$Y_s = (X_s, V(X_s), (J_{0,s}^{-1})^*). \quad (5.18)$$

It follows from (1.1) and the expression (3.2b) for the Jacobian $J_{0,t}$ that the components of Y_s solve equations of the form (3.5). One can check that both the Jacobian $\mathcal{J}_{0,s}$ for Y and its inverse $\mathcal{J}_{0,s}^{-1}$ then satisfy equations of the type (3.12). Combining these, we obtain

$$\|Z\|_{0,1,\gamma} + \|\mathcal{J}_{0,\cdot}\|_{0,1,\gamma} + \|\mathcal{J}_{0,\cdot}^{-1}\|_{0,1,\gamma} \leq K(x_0) \exp(M\|\tilde{B} + \mathcal{G}\omega\|_{0,1,\gamma}^{1/\gamma}), \quad (5.19)$$

for some constant M and some continuous function K .

Furthermore, just as in (5.4), the Malliavin derivative of Y is given by

$$\mathbf{D}_t Y_s = c_H \int_t^s \mathcal{J}_{a,s} U(Y_a) (a-t)^{H-3/2} da \mathbf{1}_{t \leq s}, \quad (5.20)$$

For some smooth function U with polynomial bounds.

We now tackle the bounds on the two very similar terms $\mathcal{I}_1, \mathcal{I}_2$. Since $t^{H-\frac{3}{2}}$ is integrable at 0 and since U grows linearly at infinity, it follows immediately that

$$\sup_{s,t \in [0,1]} |\mathbf{D}_t Y_s| \leq M \|\mathcal{J}_{0,\cdot}\|_{0,1,\infty} \|\mathcal{J}_{0,\cdot}^{-1}\|_{0,1,\infty} \|Y_{0,\cdot}\|_{0,1,\infty}. \quad (5.21)$$

Thus by (5.19) we have $\tilde{\mathbb{E}} \sup_{t \in [0,1]} \|\mathbf{D}_t v\|_{0,1,\infty}^2 < \mathbb{E}(x, \omega)$. Regarding its modulus of continuity, note that

$$\mathbf{D}_t Y_s - \mathbf{D}_t Y_r = \int_r^s \mathcal{J}_{a,s} U(Y_a) (a-t)^{H-\frac{3}{2}} da + \int_t^r (\mathcal{J}_{a,s} - \mathcal{J}_{a,r}) U(Y_a) (a-t)^{H-\frac{3}{2}} da .$$

The first term is bounded by

$$M |r-t|^{H-\frac{3}{2}} |r-s| \|\mathcal{J}_{0,\cdot}\|_{0,1,\infty} \|\mathcal{J}_{0,\cdot}^{-1}\|_{0,1,\infty} \|Y_{0,\cdot}\|_{0,1,\infty} ,$$

and the second term is bounded by

$$M |r-s|^\gamma \|\mathcal{J}_{0,\cdot}\|_\gamma \|\mathcal{J}_{0,\cdot}^{-1}\|_{0,1,\gamma} \|Y_{0,\cdot}\|_\infty ,$$

so that there does exist a constant M such that

$$|\mathbf{D}_t Y_s - \mathbf{D}_t Y_r| \leq M |r-t|^{H-\frac{3}{2}} |r-s| \|\mathcal{J}_{0,\cdot}\|_{0,1,\gamma} \|\mathcal{J}_{0,\cdot}^{-1}\|_{0,1,\gamma} \|Y_{0,\cdot}\|_\infty ,$$

uniformly over $t \leq r \leq s \leq 1$. Since $\gamma > \frac{1}{2}$, we do indeed have $\frac{3}{2} - H < 1 + \gamma - H$, so that this bound is of the type (5.16) with $\tilde{\mathbb{E}} K^p < \mathbb{E}(x, \omega)$ for every $p > 0$.

Notice furthermore that by Theorem 4.2, the matrix $C_1 = \mathcal{A}h\mathcal{A}^* \in \mathbb{R}^{n^2}$ is almost surely invertible and $\tilde{\mathbb{E}} |\mathcal{A}h\mathcal{A}^*|^{-p} < \mathbb{E}(x, \omega)$ for all $p \geq 1$. Combining this with (5.19) and the expressions for \mathcal{I}_1 and \mathcal{I}_2 , we conclude that (5.16) does indeed hold for these terms.

Let us now turn to the term \mathcal{I}_3 . Recall that the Fréchet derivative of the inverse an $n \times n$ matrix A in the direction H is given by $D^H A^{-1} = -A^{-1} H A^{-1}$. It thus follows from the chain rule that

$$\mathbf{D}_t ((\mathcal{A}h\mathcal{A}^*)^{-1} \xi) = -(\mathcal{A}h\mathcal{A}^*)^{-1} (\mathbf{D}_t (\mathcal{A}h\mathcal{A}^*)) (\mathcal{A}h\mathcal{A}^*)^{-1} \xi$$

Thus by Cauchy-Schwartz we have,

$$\left(\tilde{\mathbb{E}} \left\| \mathbf{D}_t ((\mathcal{A}h\mathcal{A}^*)^{-1} \xi) \right\|_{0,1,\infty}^p \right)^2 \leq \tilde{\mathbb{E}} |(\mathcal{A}h\mathcal{A}^*)^{-1}|^{4p} \tilde{\mathbb{E}} \left\| \mathbf{D}_t (\mathcal{A}h\mathcal{A}^*) \right\|_{0,1,\infty}^{2p} |\xi|^2 .$$

By the product rule and the definition of \mathcal{A} ,

$$\mathbf{D}_t (\mathcal{A}h\mathcal{A}^*) = \int_0^1 (\mathbf{D}_t J_{0,s}^{-1}) V(X_s) h(s) V^*(X_s) (J_{0,s}^{-1})^* ds + \dots , \quad (5.22)$$

where “...” stands for the remaining three terms that are very similar. Thus from (5.21) it follows that

$$\begin{aligned} \tilde{\mathbb{E}} \left| \int_0^1 (\mathbf{D}_t J_{0,s}^{-1}) V(X_s) h(s) V^*(X_s) (J_{0,s}^{-1})^* ds \right|^p \\ \leq M \tilde{\mathbb{E}} \|\mathcal{J}_{0,\cdot}\|_{0,1,\infty}^p \|\mathcal{J}_{0,\cdot}^{-1}\|_{0,1,\infty}^p \|Y_{\cdot}\|_{0,1,\infty}^{2p} \|V(X_{\cdot})\|_{0,1,\infty}^{2p} < \mathbb{E}(x, \omega). \end{aligned}$$

Since J^{-1} and $V(X)$ play indistinguishable roles in (5.18) and (5.19), the bounds on the remaining three terms follow in an identical fashion, so that

$$\mathbb{E} \sup_{t \in [0,1]} \left| \mathbf{D}_t ((\mathcal{A}h\mathcal{A}^*)^{-1} \xi) \right|^p < \mathbb{E}(x, \omega) .$$

Combining this with (5.19) and the definition of \mathcal{I}_3 , we finally conclude that (5.16) also holds for the term \mathcal{I}_3 (this time with $\alpha = 0$), thus concluding the proof. \square

Putting this all together, we can finally give the

Proof of Theorem 5.1. It follows from (5.6) that all we need to do is to show the bound (5.9), we have

$$\tilde{\mathbb{E}}\left(\int_0^\infty (\mathcal{D}^{H-\frac{1}{2}}v)(s) dW_s\right)^2 < \mathbb{E}(x, \omega).$$

It follows from [Nua06] that in this case one has the following extension of Itô's isometry:

$$\begin{aligned} \tilde{\mathbb{E}}\left(\int_0^\infty (\mathcal{D}^{H-\frac{1}{2}}v)(s) dW_s\right)^2 &= \tilde{\mathbb{E}}\left(\int_0^\infty |\mathcal{D}^{H-\frac{1}{2}}v(s)|^2 ds\right) \\ &\quad + \tilde{\mathbb{E}}\left(\int_0^\infty \int_0^\infty \mathbf{D}_t(\mathcal{D}^{H-\frac{1}{2}}v(s)) \mathbf{D}_s(\mathcal{D}^{H-\frac{1}{2}}v(t)) ds dt\right), \end{aligned} \quad (5.23)$$

where as before \mathbf{D} denotes Malliavin derivative with respect to the underlying Wiener process. Theorems 5.7 and 5.11 respectively show that the first and second term above are bounded above by $\mathbb{E}(x, \omega)$ and we are done. \square

6 Appendix

We give the proofs of some technical lemmas here.

Proof of lemma 2.5. The proof follows from the calculations done in [HO07]. Since the function $g(\cdot)$ defined in (2.7) is smooth everywhere except at 0, it follows from (2.6) that m_t is smooth on any interval bounded away from 0. Now a simple change of variables in (2.6) yields

$$t \frac{dm_t}{dt} = \gamma_H \int_0^\infty \frac{1}{r} t g'(t) \omega\left(-\frac{t}{r}\right) dr. \quad (6.1)$$

Since $g'(t)$ is smooth everywhere except at 0, for the map $t \mapsto t \frac{dm}{dt}(t)$ to have finite $\|\cdot\|_{0,T,\gamma}$, by (6.1) we need to only check the regularity of $t g'(t)$ at 0 and at infinity. Indeed, if we show that the function $t g'(t)$ behaves like $\mathcal{O}(t)$ for $t \ll 1$, and $\mathcal{O}(t^{H-1/2})$ for $t \gg 1$, the conclusion follows from Lemma A.1 and Proposition A.2 of [HO07]. We have,

$$\begin{aligned} t g'(t) &= t^{H-1/2} + (H-3/2)t \int_0^1 \frac{(u+t)^{H-5/2}}{(1-u)^{H-1/2}} du + M t^2 \int_0^1 \frac{(u+t)^{H-7/2}}{(1-u)^{H-1/2}} du \\ &= g(t) + M t^2 \int_0^1 \frac{(u+t)^{H-7/2}}{(1-u)^{H-1/2}} du \end{aligned} \quad (6.2)$$

for a constant M . Notice that by Lemma A.1 of [HO07] we have that $g(t)$ behaves like $\mathcal{O}(t)$ for $t \ll 1$ and $\mathcal{O}(t^{H-1/2})$ for $t \gg 1$. It can be seen easily that for $t \gg 1$, the term $t^2 \int_0^1 \frac{(u+t)^{H-7/2}}{(1-u)^{H-1/2}} du$ behaves like $\mathcal{O}(t^{H-3/2})$. For $t \ll 1$, we have

$$\begin{aligned} t^2 \int_0^1 \frac{(u+t)^{H-7/2}}{(1-u)^{H-1/2}} du &< t \int_0^1 \frac{(u+t)^{H-5/2}}{(1-u)^{H-1/2}} du \\ &< t \int_0^1 u^{H-5/2} \left(1 - (1-u)^{1/2-H}\right) du + M(t) \end{aligned} \quad (6.3)$$

where $M(t)$ is $\mathcal{O}(t)$ for $t \ll 1$. Now since $(1 - (1-u)^{1/2-H})$ is $\mathcal{O}(u)$ for $u \approx 0$, and diverges like $\mathcal{O}((1-u)^{1/2-H})$ for $u \approx 1$, the integral in (6.3) is finite. Thus from (6.2) it follows that $t g'(t)$ is $\mathcal{O}(t)$ for $t \ll 1$ and we are done. \square

Proof of Lemma 4.10. Let $f_{\min} = \inf_{t \in [0, T]} f(t)$, so we have $|f_{\min}| \leq \frac{\|f\|_{L^2[0, T]}}{\sqrt{T}}$. Therefore it follows from the definition of Hölder continuity that

$$\|f\|_{0, T, \infty} \leq \frac{\|f\|_{L^2[0, T]}}{\sqrt{T}} + \|f\|_{0, T, \gamma} T^\gamma. \quad (6.4)$$

By Chebyshev's inequality we have, for any $a > 0$, $\lambda\{x : |f(x)| > a\} \leq \frac{\|f\|_{L^2[0, T]}^2}{a^2}$ where λ denotes the Lebesgue measure. It can be seen easily that for any $a \geq |f_{\min}|$, we also have the bound

$$\|f\|_{0, T, \infty} \leq a + a^{-2\gamma} \|f\|_{0, T, \gamma} \|f\|_{L^2[0, T]}^{2\gamma}. \quad (6.5)$$

Now we combine the estimates obtained in (6.4) and (6.5). Making the choice $a = \|f\|_{0, T, \gamma}^{\frac{1}{2\gamma+1}} \|f\|_{L^2[0, T]}^{\frac{2\gamma}{2\gamma+1}}$ in (6.5) yields the bound

$$\|f\|_{0, T, \infty} \leq 2 \|f\|_{0, T, \gamma}^{\frac{1}{2\gamma+1}} \|f\|_{L^2[0, T]}^{\frac{2\gamma}{2\gamma+1}}, \quad (6.6)$$

provided that we have $\|f\|_{0, T, \gamma}^{\frac{1}{2\gamma+1}} \|f\|_{L^2[0, T]}^{\frac{2\gamma}{2\gamma+1}} \geq \frac{\|f\|_{L^2[0, T]}}{\sqrt{T}} \geq |f_{\min}|$. This is the case when $\|f\|_{0, T, \gamma} \geq \|f\|_{L^2[0, T]} T^{-\gamma-\frac{1}{2}}$. When $\|f\|_{0, T, \gamma} \leq \|f\|_{L^2[0, T]} T^{-\gamma-\frac{1}{2}}$, from (6.4) we obtain

$$\|f\|_{0, T, \infty} \leq \|f\|_{L^2[0, T]} (T^{1/4} + 1)^2, \quad (6.7)$$

thus completing the proof. \square

References

- [Arn98] L. ARNOLD. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.
- [BC07] F. BAUDOUIN and L. COUTIN. Operators associated with a stochastic differential equation driven by fractional Brownian motions. *Stochastic Process. Appl.* **117**, no. 5, (2007), 550–574.
- [BH07] F. BAUDOUIN and M. HAIRER. A version of Hörmander's theorem for the fractional Brownian motion. *Probab. Theory Related Fields* **139**, no. 3-4, (2007), 373–395.
- [Bis81] J.-M. BISMUT. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions. *Z. Wahrsch. Verw. Gebiete* **56**, no. 4, (1981), 469–505.
- [Bog98] V. I. BOGACHEV. *Gaussian measures*, vol. 62 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1998.
- [CQ02] L. COUTIN and Z. QIAN. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields* **122**, no. 1, (2002), 108–140.
- [DPEZ95] G. DA PRATO, K. D. ELWORTHY, and J. ZABCZYK. Strong Feller property for stochastic semilinear equations. *Stochastic Anal. Appl.* **13**, no. 1, (1995), 35–45.
- [DPZ96] G. DA PRATO and J. ZABCZYK. *Ergodicity for infinite-dimensional systems*, vol. 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.
- [EL94] K. D. ELWORTHY and X.-M. LI. Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125**, no. 1, (1994), 252–286.
- [Fri10] P. K. FRIZ. *Multidimensional stochastic processes as rough paths. Theory and Applications.*, vol. 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2010.

- [Hai05] M. HAIRER. Ergodicity of stochastic differential equations driven by fractional Brownian motion. *Ann. Probab.* **33**, no. 2, (2005), 703–758.
- [Hai09] M. HAIRER. Ergodic properties of a class of non-Markovian processes. In *Trends in Stochastic Analysis*, vol. 353 of *LMS Lecture Note Series*. Cambridge University Press, 2009.
- [HM06] M. HAIRER and J. C. MATTINGLY. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)* **164**, no. 3, (2006), 993–1032.
- [HM09] M. HAIRER and J. C. MATTINGLY. A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs, 2009. Preprint.
- [HN07] Y. HU and D. NUALART. Differential equations driven by Hölder continuous functions of order greater than 1/2. In *Stochastic analysis and applications*, vol. 2 of *Abel Symp.*, 399–413. Springer, Berlin, 2007.
- [HO07] M. HAIRER and A. OHASHI. Ergodic theory for SDEs with extrinsic memory. *Ann. Probab.* **35**, no. 5, (2007), 1950–1977.
- [Hör67] L. HÖRMANDER. Hypoelliptic second order differential equations. *Acta Math.* **119**, (1967), 147–171.
- [KS84] S. KUSUOKA and D. STROOCK. Applications of the Malliavin calculus. I. In *Stochastic analysis (Katata/Kyoto, 1982)*, vol. 32 of *North-Holland Math. Library*, 271–306. North-Holland, Amsterdam, 1984.
- [KS85] S. KUSUOKA and D. STROOCK. Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32**, no. 1, (1985), 1–76.
- [KS87] S. KUSUOKA and D. STROOCK. Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34**, no. 2, (1987), 391–442.
- [LQ02] T. LYONS and Z. QIAN. *System control and rough paths*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2002. Oxford Science Publications.
- [Mal78] P. MALLIAVIN. Stochastic calculus of variations and hypoelliptic operators. *Symp. on Stoch. Diff. Equations, Kyoto 1976* 147–171.
- [Mal97] P. MALLIAVIN. *Stochastic analysis*, vol. 313 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997.
- [Mis08] Y. S. MISHURA. *Stochastic calculus for fractional Brownian motion and related processes*, vol. 1929 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008.
- [MN08] L. MAZLIAK and I. NOURDIN. Optimal control for rough differential equations. *Stoch. Dyn.* **8**, no. 1, (2008), 23–33.
- [MT93] S. P. MEYN and R. L. TWEEDIE. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.
- [MVN68] B. B. MANDELBROT and J. W. VAN NESS. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10**, (1968), 422–437.
- [Nor86] J. NORRIS. Simplified Malliavin calculus. In *Séminaire de Probabilités, XX, 1984/85*, vol. 1204 of *Lecture Notes in Math.*, 101–130. Springer, Berlin, 1986.
- [NS06] I. NOURDIN and T. SIMON. On the absolute continuity of one-dimensional SDEs driven by a fractional Brownian motion. *Statist. Probab. Lett.* **76**, no. 9, (2006), 907–912.
- [NS09] D. NUALART and B. SAUSSEREAU. Malliavin calculus for stochastic differential equations driven by a fractional Brownian motion. *Stochastic Process. Appl.* **119**, no. 2, (2009), 391–409.

- [Nua06] D. NUALART. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second ed., 2006.
- [SKM93] S. G. SAMKO, A. A. KILBAS, and O. I. MARICHEV. *Fractional integrals and derivatives*. Gordon and Breach Science Publishers, Yverdon, 1993. Theory and applications, Edited and with a foreword by S. M. Nikol'skii, Translated from the 1987 Russian original, Revised by the authors.
- [ST94] G. SAMORODNITSKY and M. S. TAQQU. *Stable non-Gaussian random processes*. Stochastic Modeling. Chapman & Hall, New York, 1994. Stochastic models with infinite variance.
- [You36] L. C. YOUNG. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* **67**, no. 1, (1936), 251–282.